CORRECTED
VERSION*

RU-0073

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68, C12P 19/34, C07K 13/00, G06F 17/00	**A1**	(11) International Publication Number: **WO 00/05414**
		(43) International Publication Date: 3 February 2000 (03.02.00)

(72) Inventors; and
(75) Inventors/Applicants (for US only): ANDERSON, Stephen [US/US]; 158 Springdale Road, Princeton, NJ 08540 (US). MONTELIONE, Gaetano [US/US]; 127 North 5th Avenue, Highland Park, NJ 08904 (US). HUANG, Yuanpeng [US/US]; ASB, Annex 11, 58 Bevier Road, Piscataway, NJ 08854–8010 (US).

(54) Title: LINKING GENE SEQUENCE TO GENE FUNCTION BY THREE DIMENSIONAL (3D) PROTEIN STRUCTURE DETERMINATION

(57) Abstract

The present invention provides a structure–functional analysis engine for the high–throughput determination of the biochemical function of protein domains of unknown function, as exemplified in the flowchart of the figure. The present invention uses bioinformatics, molecular biology and nuclear magnetic resonance tools for the rapid and automated determination of the three–dimensional structures of proteins and protein domains.

NOESY_ASSIGN PROCESS

STEP 1: SIMPLE MATCH—GET ALL POSSIBLE ASSIGNMENT (A–TYPE MATCH).

STEP 2: IDENTIFY EXPECTED PEAKS WHICH ARE INTRA/SEQ. OR CONSISTENT WITH SECONDARY STRUCTURE. PUT IN U AND E SET. CREATE CURV.

STEP 3: DEFINE LOCAL MATCH TOLERANCE FOR HX DIMENSION BASED ON ASSIGNED HX RESONANCES FROM E SET.

STEP 4: SUPPLEMENT U BASED ON CHEMICAL SHIFT (UNAMBIGUOUS) WITH NOISE FILTER.

STEP 5: DEFINE B–TYPE MATCHES OF A–TYPE MATCHES.

STEP 6: USE CF OF E TO ASSIGN B TO U.

STEP 7: USE SYM TO ASSIGN B TO U.

STEP 8: USE HP–CORE TO ASSIGN B TO U.

STEP 9: COMPUTE 3D STRUCTURE.

STEP 10: VALIDATION—REMOVE FROM U LIST THE CONSISTENTLY VIOLATED NOE ASSIGNMENTS.

STEP 11: IDENTIFY EXPECTED PEAKS THAT ARE CONSISTENT WITH 3D STRUCTURE. PUT IN U SET.

STEP 12: BACK CALCULATE ALL POSSIBLE NOE'S THAT ARE EXPECTED FROM STRUCTURE. OUTPUT ANY PREDICTED ASSIGNMENTS NOT IN U OR A LIST AND PEAKS STILL IN THE A LIST.